Automated identification of an aspirin-exacerbated respiratory disease cohort

Cahill, Katherine N ; Johns, Christina B ; Cui, Jing ; Wickner, Paige ; Bates, David W ; Laidlaw, Tanya M ; Beeler, Patrick E

Abstract: BACKGROUND Aspirin-exacerbated respiratory disease (AERD) is characterized by 3 clinical features: asthma, nasal polyposis, and respiratory reactions to cyclooxygenase-1 inhibitors (nonsteroidal anti-inflammatory drugs). Electronic health records (EHRs) contain information on each feature of this triad. OBJECTIVE We sought to determine whether an informatics algorithm applied to the EHR could electronically identify patients with AERD. METHODS We developed an informatics algorithm to search the EHRs of patients aged 18 years and older from the Partners Healthcare system over a 10-year period (2004-2014). Charts with search terms for asthma, nasal polyps, and record of respiratory (cohort A) or unspecified (cohort B) reactions to nonsteroidal anti-inflammatory drugs were identified as "possible AERD." Two clinical experts reviewed all charts to confirm a diagnosis of "clinical AERD" and classify cases as "diagnosed AERD" or "undiagnosed AERD" on the basis of physician-documented AERD-specific terms in patient notes. RESULTS Our algorithm identified 731 "possible AERD" cases, of which 638 were not in our AERD patient registry. Chart review of cohorts A (n = 511) and B (n = 127) demonstrated a positive predictive value of 78.4% for "clinical AERD," which rose to 88.7% when unspecified reactions were excluded. Of those with clinical AERD, 12.4% had no mention of AERD by any treating caregiver and were classified as "undiagnosed AERD." "Undiagnosed AERD" cases were less likely than "diagnosed AERD" cases to have been seen by an allergist/immunologist (38.7% vs 93.2%; P < .0001). CONCLUSIONS An informatics algorithm can successfully identify both known and previously undiagnosed cases of AERD with a high positive predictive value. Involvement of an allergist/immunologist significantly increases the likelihood of an AERD diagnosis.

DOI: https://doi.org/10.1016/j.jaci.2016.05.048

DOI: https://doi.org/10.1016/j.jaci.2016.05.048

Title: Automated identification of an aspirin-exacerbated respiratory disease cohort

List of Authors:

Katherine N. Cahill, MD[1,2], Christina B. Johns, BA[2], Jing Cui, MD, PhD[1,2], Paige Wickner, MD, MPH[1,2], David W. Bates MD, MSc[1,3], Tanya M. Laidlaw, MD[1,2] and Patrick E. Beeler, MD[1,3,4].

[1]Department of Medicine, Harvard Medical School, Boston, MA, USA; [2]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, MA, USA; [3]Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA; [4]Research Center for Medical Informatics, University Hospital Zurich and University of Zurich, Zurich, Switzerland.

Corresponding author:

Katherine N. Cahill, MD

Brigham and Women's Hospital

1 Jimmy Fund Way, Smith Building, Room 626B

Boston, MA 02115

kncahill@partners.org

Tel: 617-525-1300, Fax: 617-525-1310

Funding sources:

This work was supported by NIH grants NIH K23 HL111113 and NIH K23 AI118804-01, the Swiss National Science Foundation (Dr. Beeler), departmental funds and by generous contributions from the Vinik and Kaye Families.

Abstract:

Background: Aspirin-exacerbated respiratory disease (AERD) is characterized by three clinical features: asthma, nasal polyposis, and respiratory reactions to cyclooxygenase-1 inhibitors (NSAIDs). Electronic health records (EHRs) contain information on each feature of this triad.

Objective: To determine if an informatics algorithm applied to the EHR could electronically identify patients with AERD.

Methods: We developed an informatics algorithm to search the EHRs of patients age 18 and older from the Partners Healthcare system over a 10 year period (2004-2014). Charts with search terms for asthma, nasal polyps and record of respiratory (Cohort A) or unspecified (Cohort B) reactions to NSAIDs were identified as "possible AERD". Two clinical experts reviewed all charts to confirm a diagnosis of "clinical AERD" and classify cases as "diagnosed AERD" or "undiagnosed AERD" based on physician documented AERD-specific terms in patient notes.

Results: Our algorithm identified 731 "possible AERD" cases, of which 638 were not in our AERD patient registry. Chart review of cohorts A (n=511) and B (n=127) demonstrated a positive predictive value (PPV) of 78.4% for "clinical AERD", which rose to 88.7% when unspecified reactions were excluded. Of those with clinical AERD, 12.4% had no mention of AERD by any treating caregiver and were classified as "undiagnosed AERD". "Undiagnosed AERD" cases were less likely to have been seen by an allergist/immunologist than "diagnosed AERD" cases (38.7% vs. 93.2%, P<.0001).

Conclusion: An informatics algorithm can successfully identify both known and previously undiagnosed cases of AERD with a high PPV. Involvement of an allergist/immunologist significantly increases the likelihood of an AERD diagnosis.

Key Messages:

- An informatics algorithm can be used to search electronic health records to identify diagnosed and previously undiagnosed cases of clinical aspirin-exacerbated respiratory disease (AERD).
- Incomplete recording of drug reaction data by caregivers limits the PPV of this algorithm.
- Involvement of allergy/immunology specialists in the care of subjects with asthma, nasal polyposis, and NSAID allergy increases the likelihood that a diagnosis of AERD will be made.

Capsule Summary:

An informatics search algorithm can successfully identify diagnosed and undiagnosed cases of aspirin-exacerbated respiratory disease (AERD) in the electronic health record.

Key Words:

Aspirin-exacerbated respiratory disease

Electronic health record

Asthma

Nasal polyps

Non-steroidal anti-inflammatory drugs

Chronic rhinosinusitis

Structured query language

Clinical decision support

Abbreviations:

Aspirin-exacerbated respiratory disease (AERD)

Electronic health record (EHR)

Non-steroidal anti-inflammatory drugs (NSAIDs)

Cyclooxygenase-1 (COX-1)

Chronic rhinosinusitis with nasal polyposis (CRSwNP)

International classification of diseases 9 (ICD-9)

Partners Research Patient Data Repository (RPDR)

Structured query language (SQL)

Positive predictive value (PPV)

Interquartile range (IQR)

Confidence interval (CI)

Introduction:

Electronic health records (EHR) provide the advantage of an electronically searchable patient chart and are now being widely used in North American and Europe. One of the ways EHRs can be used to improve patient care is to develop informatics algorithms for disease diagnosis. Using this approach, cohorts of patients with disease-specific characteristics can be identified for diagnosis (1). Identified patients may then benefit in a variety of ways, such as from disease-targeted therapeutics and from participation in clinical trials and translational research investigations. This may be particularly important in the field of clinical allergy and immunology where many of the common diseases encountered lack accurate disease-specific coding in our current systems.

In the classic triad form, aspirin-exacerbated respiratory disease (AERD), also referred to as Samter's Triad, is the unique clinical combination of chronic rhinosinusitis with nasal polyposis (CRSwNP), asthma, and respiratory reactions to all inhibitors of cyclooxygenase (COX)-1. The syndrome affects 7.2% (95% CI, 5.26% to 9.03%) of adults with asthma and 14.9% (95% CI, 6.48% to 23.29%) of those with severe asthma, and therefore may affect up to 2 million U.S. adults (2). Ingestion of aspirin or any COX-1 inhibitor elicits hypersensitivity reactions within 30 minutes to 3 hours that include worsening upper respiratory symptoms and acute bronchoconstriction, sometimes requiring emergency medical care. Although there are patients with respiratory reactions to COX-1 inhibitors who do not have all three components of this disease (3-5), we will consider the classic triad for the duration of this manuscript. AERD is a chronic medical condition that dramatically impacts quality of life and medical resource utilization beyond that of most aspirin-tolerant patients with asthma or CRSwNP (6). Despite the morbidity of the syndrome and its frequency in the adult asthmatic population, our clinical

experience is that there is a delay of many months to years between the onset of AERD symptoms and a formal diagnosis (4), and research efforts in AERD are hampered by modest sample sizes.

Unfortunately, AERD lacks a unifying ICD-9 or ICD-10 code. Since AERD is characterized by a unique triad, we hypothesized that the simultaneous use of ICD-9 codes for asthma and nasal polyps, problem list entries, and medication allergy entries would automatically identify a cohort of possible AERD cases. Therefore, we developed and tested an EHR algorithm to identify subjects with AERD.

Methods:

Informatics Algorithms

Applying an informatics algorithm to the Partners Research Patient Data Repository (RPDR) (1, 7), the EHRs at 2 academic hospitals (Massachusetts General Hospital and Brigham and Women's Hospital [BWH]) and one community hospital (Faulkner Hospital) affiliated with the Partners Healthcare system were searched over a 10 year period (12/2004-11/2014). IRB approval was obtained for this study. The EHR at the institutions searched is entirely electronic and included both inpatient and outpatient data from any affiliated hospital or clinic. All charts of patients age 18 or older who had one or more encounters during this time period were searched for AERD-relevant features. One RPDR query (Repository Table E1) was designed to find patients with ICD-9 codes, problem list entries, laboratory values (eosinophils >500/μL) or medications associated with asthma *and* with ICD-9 codes, problem list entries, intranasal steroids or surgical billing codes related to nasal polyposis. The second RPDR query was

designed to find patients with NSAID allergy. The union of the two RPDR queries resulted in datasets including 168,126 patients, which were further processed as described below.

The datasets obtained from RPDR were preprocessed, i.e. decrypted and decompressed, and aggregation algorithms were used to summarize the resulting raw data tables, enabling first reviews of the data. Because the RPDR queries were designed to capture all patients of potential interest, structured query language (SQL) statements were used to filter and analyze patient data and allow for the identification of the most important structured terms used in the final algorithm.

Three preliminary SQL queries were developed for each characteristic of AERD, searching the data tables for specific terms, e.g. “asthma”, and misspellings such as “amaphylaxis” were also considered. Each query returned one patient population with asthma (Repository Figure E1), one population with nasal polyps (Repository Figure E2) and one population with NSAID allergy (Repository Figure E3). The NSAID allergy SQL was designed to identify charts that reported reactions typical of the respiratory symptoms triggered by NSAIDs in AERD or charts that reported unspecified (“unknown”) reactions to NSAIDs. Reaction types not classically associated with AERD, e.g. gastritis or urticaria, were excluded. The results (patient sets) of each query were used to further refine the SQL queries filtering more specific data about the identified populations. The BWH AERD patient registry (n=96), a well-phenotyped database of patients with aspirin-challenge confirmed AERD, was also used to identify information of increased significance, and the SQL queries were iteratively revised several times. In the example of nasal polyps, if a problem was noted by a clinician that did not contain the necessary key words but one of the terms “sinus”, “nasal” or “allergic rhinitis”, then the problem-associated comment was searched for “polyp”.

Over the course of these iterations, it became clear that *diagnoses* (ICD-9 codes), *problems* including associated *comments*, and *allergens,* focusing only on those with specified respiratory (e.g. bronchospasm, wheeze, nasal congestion) reactions, or unspecified reactions to any inhibitor of COX-1*,* were the most important components to identify potential AERD patients. The intersection of the three populations identified "possible AERD" cases (Figure 1), which were further stratified by the type of reaction to an NSAID recorded in the EMR; Cohort A included cases where specific respiratory symptoms were recorded and Cohort B included cases where the reaction symptoms were unspecified, i.e. recorded as "unknown".

A number of cases identified as "possible AERD" were already recorded as having known AERD within structured information in the EHR e.g. problem lists and allergies and/or through involvement in the BWH AERD patient registry. Therefore a fourth SQL query (Supplementary File E5) was set up that searched only for AERD-specific terms within structured information in the EHR, to determine if that more simplified approach would be sufficient to identify cases of AERD from the EHR.

Chart Reviews

Two allergy/immunology experts with a clinical focus on AERD independently performed chart reviews. All charts from Cohort A and Cohort B were reviewed by at least one reviewer, with 20 charts from each cohort reviewed by both reviewers to assess the inter-rater agreement (Kappa). Reviewers defined "clinical AERD" as the presence of an asthma diagnosis, nasal polyps and a report of a classical respiratory reaction to one or more NSAIDs. The presence of nasal polyposis was confirmed during chart review if one of the following criteria were met: 1) documentation of rhinoscopic evidence of nasal polyposis, 2) surigical/pathologic

report confirming nasal polyposis, or 3) radiologic evidence of nasal polyposis. Cases which carried a diagnosis of cystic fibrosis, sinus malignancy, or unilateral sinus disease or were determined by chart review to either not meet criteria for a diagnosis of AERD or not have sufficient information recorded within their chart to determine the diagnosis, were labeled "Not AERD". During this review, unstructured EHR data, including progress, hospital visit, and surgical procedure notes, were reviewed using a queriable patient inference dossier (8) to identify if a caregiver had made a prior diagnosis of AERD (or another term for the disease, including Samter's triad, aspirin-sensitive asthma, aspirin-intolerant asthma, or triad asthma) that was not recorded in the structured data. These cases were defined as "diagnosed AERD." Cases established by expert review as having "clinical AERD" but lacking any documentation of AERD in either the structured or unstructured data within the EHR were considered "undiagnosed AERD." Whether the patient had ever had clinical involvement of pulmonary, allergy/immunology and otolaryngology specialists in each case was noted.

Statistical Analyses

All data are represented as mean ± standard deviation (SD) unless otherwise noted. Cohen's kappa coefficient was used to measure inter-rater agreement on the clinical diagnosis of AERD by our expert reviewers. Positive-predictive values (PPV) were calculated from chart reviews of Cohort A, B, and the BWH AERD registry. Fisher's exact test was used to assess differences in gender and race between "diagnosed" and "undiagnosed" AERD; a Mann-Whitney U test was employed to determine difference in age. Differences in rates of specialty physician evaluations were assessed using a contingency table and Fisher's exact test. T tests were performed to determine differences in number of encounters. GraphPad Prism version 6.07 for windows, GraphPad Software, La Jolla California USA, www.graphpad.com, SAS software,

version 9.4, Cary, North Carolina, USA, and/or R version 3.2.1, R Foundation for Statistical Computing, Vienna, Austria was used to complete these analyses.

Results

A total of 2,647,842 charts were queried using RPDR between 12/1/2004 and 11/30/2014. The cohort defined by the intersection of the asthma (SQL #1), nasal polyp (SQL #2) and NSAID allergy (SQL #3) queries was considered to contain "possible AERD" cases (n=732, Figure 1). One case was identified as a test patient, a virtual patient generated for training purposes, and excluded and 93 cases participated in the AERD registry and had known confirmed AERD. Of the remaining 638 cases, Cohort A (n=511) included cases with record of a respiratory reaction to NSAIDs and Cohort B (n=127) included cases with an unspecified reaction to an NSAID (Figure 2).

Cohorts A and B were independently reviewed by both reviewers. The inter-rater agreement value, kappa, for each cohort was 100%. The PPV for the identification of "clinical AERD" cases using this informatics algorithm is 81.1% (Cohorts A, B, and the BWH AERD Registry). The PPV excluding the AERD registry charts (Cohort A and B) is 78.4% which rises to 88.7% if only cases with a specified respiratory reaction to an NSAID (Cohort A) are considered. After expert review of progress notes, 12.4% of "clinical AERD" cases identified (11.9% in Cohort A and 17.0% in Cohort B) were labeled "undiagnosed AERD", indicating that the expert review agreed they had the triad of clinical symptoms consistent with AERD but there was no mention of AERD or a similar term in the EHR (Figure 2). Significantly less involvement from allergy/immunology specialists was noted in the care of "undiagnosed AERD" cases as compared with "diagnosed AERD" cases (38.7% vs. 93.2%, P<..0001; Table I). Among those

"clinical AERD" patients who had been evaluated by only one type of specialty provider, 100% of the 6 cases seen by only allergy, 40.9% of the 44 cases seen by only ENT, and 33.3% of the 3 cases seen by only pulmonary were recorded in the EHR as having been diagnosed with AERD (P<.05).

The patient demographics of "diagnosed" and "undiagnosed AERD" and the BWH AERD patient registry are reported in Table II. The diagnosed AERD cohort median age (interquartile range (IQR)) was slightly younger than the undiagnosed cohort (54 (IQR=45-65), 58 (IQR=51-72), respectively, P<.01). There was no significant difference in sex or race between cohorts. The median number of patient encounters with the Partners Healthcare system was not different between those with "diagnosed" and "undiagnosed" AERD (37.5 (IQR=11-101) and 54.5 (IQR=19-126), respectively, P=.31).

Application of SQL #4 (n=255) identified only 28.9% (n=211) of the "possible AERD" cases (Figure 3) and an additional 44 cases not identified by the EHR algorithm. Of the 42 charts in SQL#4 not identified by the EHR algorithm or included in the BWH AERD registry, 20 lacked one or more components of the triad and were considered "Not AERD" and 22 (52.4%) were labeled "clinical AERD" after expert chart review. Application of the primary EHR search algorithm to just the BWH AERD patient registry identified 93 of 96 patients (96.9%). Of the 3 cases from the BWH AERD patient registry that were not identified by the AERD algorithm, 2 had no NSAID allergy recorded, representing serious omissions that impact patient safety, and one lacked appropriate documentation of nasal polyps. Taken together, our primary algorithm failed to identify 3.7% [23 of 618 (Clinical AERD (n=500) + BWH AERD Registry (n=96) + SQL#4 Clinical AERD (n=22)] of the known patients with AERD in the EHR.

Discussion

We demonstrate that an informatics algorithmic approach can be used to identify both diagnosed and previously undiagnosed cases of AERD. Our approach identified 593 known or expert-confirmed cases of AERD with a PPV of 81.1% while missing only 3.7% of the known patients with AERD in the EHR. Among those cases identified by our algorithm and confirmed by expert review as having "clinical AERD", 12.4% (n=62) carried no mention of AERD or an equivalent term in the medical chart. As far as could be determined from their medical chart, no caregiver had ever realized the connection between their clinical triad of symptoms and therefore these cases had never been given the diagnosis of AERD (Figure 2). Patients in this "undiagnosed AERD" category were less likely to have been evaluated by an allergy/immunology specialist (Table I), highlighting the role of allergist/immunologists in correctly identifying this disease. Cases of "undiagnosed AERD" identified by the algorithm have not yet been exposed to the gold standard for diagnosis of AERD, aspirin challenge, to confirm the assessment made by our expert clinicians. The current literature suggests up to 15% of those cases meeting clinical criteria for AERD may have a negative aspirin challenge (9, 10). However, the clinical experience from our institution involving more than 150 aspirin challenges is that <5% of patients with asthma, nasal polyposis and a historical respiratory reaction to an NSAID go on to have a negative aspirin challenge (4). This suggests our informatics algorithm can identify new diagnoses of AERD and could facilitate access to disease-specific treatments for these patients, which have been shown to improve their care (11-13).

Algorithm-identified cases of AERD, both "diagnosed" and "undiagnosed", demonstrate the classical female predominance (9, 10, 14). The slightly younger age in the "diagnosed AERD" cases cannot easily be explained with the data generated in this study (Table II). One

hypothesis drawn from our clinical experience is that younger patients with AERD are using electronic resources to connect their triad of symptoms and may present to their providers questioning a diagnosis of AERD, leading to greater consideration and confirmation of AERD. Previously there has been no racial predilection for the development of AERD reported and our racial demographics reflect the racial distribution of the Partners Healthcare patient population. Race does not predict if a case is diagnosed or not. No data about asthma severity/control was collected/analyzed and no conclusions can be made about the nature of the upper or lower respiratory disease in the cohorts. The lack of a difference in number of encounters between the groups suggests both groups utilize the healthcare system at similar rates, had similar amounts of data available for chart review, and that the number of encounters with the healthcare system did not bias towards identifying an "undiagnosed" case of AERD.

The benefit of using such an algorithm to identify patients with AERD is multi-factorial. In the short-term, patients with AERD would have better access to disease-specific therapy including zileuton which improves nasal symptoms and FEV1(15) and high-dose aspirin therapy which improves sinus and asthma symptom scores and decreases nasal congestion, corticosteroid use (oral and inhaled), the number of sinus infections per year, and the need for repeat polypectomy (16, 17). Additionally, of the cases of AERD we identified, less than 20% are participating in the BWH AERD patient registry. As patients who participate in the registry are provided with formal educational materials about their disease and are offered involvement in research opportunities, this highlights the potential to engage 500 new subjects in clinical or translational research focused on AERD at our or any other institution. Use of an informatics algorithm at any institution employing an electronic medical record to identify patients with AERD, a disease lacking a unifying ICD-9 or 10 code or diagnostic laboratory test, has the

power to improve patient care immediately and to support the research endeavors that will yield future advances in patient care.

The algorithm we present used commonly-coded information for diagnosis, billing and allergy information that is captured in any electronic medical record. Our development of an EHR-based phenotyping algorithm for AERD can be deployed in other electronic medical record programs, both nationally and internationally, which are capturing data on the diagnosis of asthma, nasal polyps and allergy to NSAIDs (18). Similar algorithms for rheumatoid arthritis, drug-induced liver injury and genomic phenotyping have been successfully employed across 2-13 different EHR platforms (19-21). The data model employed by our EHR does not differ substantially from other EHRs both nationally and internationally. Minor adjustments for language and regional differences in terminology (e.g. NERD, i.e. NSAID-exacerbated respiratory disease, which is commonly used in Europe) would be required to maximize the success of adapting this algorithm. Although we have generated this algorithm and searched the patient charts from two large referral-based tertiary care centers with active research programs in asthma, nasal polyps, and AERD, the data used to identify potential cases of AERD is basic information that should be captured by primary care and specialist providers even if they have no knowledge of AERD.

As with all informatics algorithms, our algorithm is limited by the amount and the quality of the data contained within the EHR, specifically among the details of drug allergy recordings (22). The PPV of our algorithm drops from 88.7% (Cohort A) to 78.4% (Cohort A and B) if we include cases in which the symptoms of reaction to NSAIDs are not specified. Of those cases in Cohort A determined not to have AERD, 21 of 58 of them were classified as such because they lacked a sufficient NSAID allergy history in the chart to meet our pre-specified criteria for

characterization as AERD. The inclusion of SQL#4 confirms that use of AERD-specific search terms alone vastly underestimates the potential cases of AERD in the EHR (Figure 3). A closer look at those 42 charts identified by SQL #4 which were not found by the primary EHR AERD algorithm or included in the BWH AERD registry highlights the danger of incomplete and inaccurate information contained within the EHR. 47.6% (n=20) of these charts were eventually classified as "Not AERD" due to one of two reasons: 1) AERD had initially been considered and/or recorded by a provider but then ruled out by a negative aspirin challenge or 2) the EHR did not have enough information to confirm a diagnosis of AERD. Because of these data quality limitations, use of any algorithm is likely to under-detect possible cases of AERD and no conclusions about the prevalence of AERD can be drawn from this study. In primary care settings, relying on a single ICD-9 or 10 code for the diagnosis of asthma lacks specificity (23). The requirement for multiple ICD-9/10 codes and/or additional data, e.g. concomitant prescriptions for disease-targeted therapy such as β-agonists, may be necessary to improve the specificity of this algorithm. However, no improvement in the algorithm methods can make up for the omission of information in the EHR. Our work underscores the need for complete and specific data entry in the EHR in order to maximize the patient safety and research potential.

In our healthcare system with more than 2,000,000 patient records between 11/2004-11/2014, given the known prevalence of asthma in US adults is 7% (24), and the prevalence of AERD is estimated at 7% of adults with asthma (2), we would have predicted to find >10,000 cases of AERD. In addition to the data quality issues our algorithm identified, patients referred from an outside provider to a tertiary care center for specialty care may lack complete EHR data, specifically ICD-9 coding or problem list entries for asthma or nasal polyps, if those problems are not being addressed by the specialty provider. We focused our efforts on the identification of

the classic triad of AERD (25), and did not focus on identifying those cases which lack either asthma or nasal polyposis but demonstrate the sterotypical respiratory reaction following the ingestion of a COX-1 inhibitor (3-5), likely missing these non-classic presentations of AERD. Additionally, our hospital system is known for oncology, rheumatology, and obstetric care and our starting population likely is over represented for these conditions which do not have any association with asthma.

The patient population searched presents two unique characteristics about the charts queried. First, the tertiary care setting may result in incomplete health records, as discussed above, and bias the algorithm and the chart review against assigning a diagnosis of AERD. Given the lack of disease-unique therapeutics or laboratory values in AERD, no other recorded data points can be depended upon to adequately replace missing diagnoses. Second, our cohort is likely to have more AERD-specific information available within the EHR, specifically in the problem list where an "aspirin-intolerant asthma with nasal polyposis" problem has been created at the request of BWH AERD Center physicians. We anticipate higher rates of "undiagnosed AERD" would be identified by application of this algorithm to another setting that does not have an active AERD clinical and research program. The algorithm we present does not require an AERD-specific term, which SQL#4 demonstrated was neither sensitive nor specific for AERD, and application of this algorithm approach to another EHR should have no impact on the clinically significant identification of cases which fall into Cohorts A and B.

New strategies employing the EHR to increase identification of patients with AERD and other allergic diseases hold great promise for improving clinical care and expanding access to specialists in the field. A recent survey of subjects with AERD highlighted the disconnect between beneficial therapies and their use in patients with AERD. 91% of AERD subjects

reported aspirin therapy was beneficial but <50% of the survey population had been offered aspirin therapy (6). The present algorithmic approach could be used to display automatic alert notifications to physicians in order to promote the consideration of AERD and improve documentation of AERD (26), while offering evidence-based information and detailed advice including referral options (27). Providing patients with an accurate diagnosis may empower them to seek out effective treatments for their disease and/or engage in clinical trials that have the potential to transform the future of AERD-specific care. The high PPV of our algorithm would likely generate notifications at low risk for inducing alert fatigue (28). In addition, this algorithm could be used to prioritize the generation of medication alerts for NSAID prescriptions in those patients who have a record of NSAID allergy in conjunction with a history of asthma and/or nasal polyps (29). Future work assessing the gains in patient care and safety from such an approach is needed.

AERD is an under-recognized but important disease in which current technology can be employed to better serve the needs of our patients. Leveraging the power of the EHR to identify new diagnoses has the potential to shorten the length of time between symptom onset and diagnosis and to positively affect care for patients with AERD.

References

1. Nigwekar SU, Solid CA, Ankers E, Malhotra R, Eggert W, Turchin A, et al. Quantifying a rare disease in administrative data: the example of calciphylaxis. J Gen Intern Med. 2014;29 Suppl 3:S724-31.
2. Rajan JP, Wineinger NE, Stevenson DD, White AA. Prevalence of aspirin-exacerbated respiratory disease among asthmatic patients: A meta-analysis of the literature. The Journal of allergy and clinical immunology. 2015;135(3):676-81 e1.
3. Kowalski ML, Asero R, Bavbek S, Blanca M, Blanca-Lopez N, Bochenek G, et al. Classification and practical approach to the diagnosis and management of hypersensitivity to nonsteroidal anti-inflammatory drugs. Allergy. 2013;68(10):1219-32.
4. Lee-Sarwar K, Johns C, Laidlaw TM, Cahill KN. Tolerance of daily low-dose aspirin does not preclude aspirin-exacerbated respiratory disease. The journal of allergy and clinical immunology In practice. 2015;3(3):449-51.
5. Bochenek G, Kuschill-Dziurda J, Szafraniec K, Plutecka H, Szczeklik A, Nizankowska-Mogilnicka E. Certain subphenotypes of aspirin-exacerbated respiratory disease distinguished by latent class analysis. The Journal of allergy and clinical immunology. 2014;133(1):98-103 e1-6.
6. Ta V, White AA. Survey-Defined Patient Experiences With Aspirin-Exacerbated Respiratory Disease. The journal of allergy and clinical immunology In practice. 2015.
7. Nalichowski R, Keogh D, Chueh HC, Murphy SN. Calculating the benefits of a Research Patient Data Repository. AMIA Annu Symp Proc. 2006:1044.
8. Zalis M, Harris M. Advanced search of the electronic medical record: augmenting safety and efficiency in radiology. J Am Coll Radiol. 2010;7(8):625-33.
9. White AA, Stevenson DD, Simon RA. The blocking effect of essential controller medications during aspirin challenges in patients with aspirin-exacerbated respiratory disease. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology. 2005;95(4):330-5.
10. Pleskow WW, Stevenson DD, Mathison DA, Simon RA, Schatz M, Zeiger RS. Aspirin-sensitive rhinosinusitis/asthma: spectrum of adverse reactions to aspirin. The Journal of allergy and clinical immunology. 1983;71(6):574-9.
11. Shaker M, Lobb A, Jenkins P, O'Rourke D, Takemoto SK, Sheth S, et al. An economic analysis of aspirin desensitization in aspirin-exacerbated respiratory disease. The Journal of allergy and clinical immunology. 2008;121(1):81-7.
12. Ibrahim C, Singh K, Tsai G, Huang D, Mazza J, Rotenberg B, et al. A retrospective study of the clinical benefit from acetylsalicylic acid desensitization in patients with nasal polyposis and asthma. Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology. 2014;10(1):64.
13. Bochenek G, Szafraniec K, Kuschill-Dziurda J, Nizankowska-Mogilnicka E. Factors associated with asthma control in patients with aspirin-exacerbated respiratory disease. Respiratory medicine. 2015;109(5):588-95.
14. Stevenson DD, Szczeklik A. Clinical and pathologic perspectives on aspirin sensitivity and asthma. The Journal of allergy and clinical immunology. 2006;118(4):773-86; quiz 87-8.
15. Dahlen B, Nizankowska E, Szczeklik A, Zetterstrom O, Bochenek G, Kumlin M, et al. Benefits from adding the 5-lipoxygenase inhibitor zileuton to conventional therapy in aspirin-intolerant asthmatics. Am J Respir Crit Care Med. 1998;157(4 Pt 1):1187-94.
16. Swierczynska-Krepa M, Sanak M, Bochenek G, Strek P, Cmiel A, Gielicz A, et al. Aspirin desensitization in patients with aspirin-induced and aspirin-tolerant asthma: a double-blind study. The Journal of allergy and clinical immunology. 2014;134(4):883-90.

17. Stevenson DD, Hankammer MA, Mathison DA, Christiansen SC, Simon RA. Aspirin desensitization treatment of aspirin-sensitive patients with rhinosinusitis-asthma: long-term outcomes. The Journal of allergy and clinical immunology. 1996;98(4):751-8.
18. Cho I, Lee JH, Choi SK, Choi JW, Hwang H, Bates DW. Acceptability and feasibility of the Leapfrog computerized physician order entry evaluation tool for hospitals outside the United States. Int J Med Inform. 2015;84(9):694-701.
19. Carroll RJ, Thompson WK, Eyler AE, Mandelin AM, Cai T, Zink RM, et al. Portability of an algorithm to identify rheumatoid arthritis in electronic health records. J Am Med Inform Assoc. 2012;19(e1):e162-9.
20. Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, Murphy S, et al. A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury. J Am Med Inform Assoc. 2013;20(e2):e243-52.
21. Newton KM, Peissig PL, Kho AN, Bielinski SJ, Berg RL, Choudhary V, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. J Am Med Inform Assoc. 2013;20(e1):e147-54.
22. Zimmerman CR, Chaffee BW, Lazarou J, Gingrich CA, Russell CL, Galbraith M, et al. Maintaining the enterprisewide continuity and interoperability of patient allergy data. Am J Health Syst Pharm. 2009;66(7):671-9.
23. Xi N, Wallace R, Agarwal G, Chan D, Gershon A, Gupta S. Identifying patients with asthma in primary care electronic medical record systems Chart analysis-based electronic algorithm validation study. Can Fam Physician. 2015;61(10):e474-83.
24. Prevention CfDCa. Asthma: Data, Statistics, and Surveillance 2013 [cited 2016 February 12, 2016]. Available from: http://www.cdc.gov/asthma/asthmadata.htm.
25. Berges-Gimeno MP, Simon RA, Stevenson DD. The natural history and clinical characteristics of aspirin-exacerbated respiratory disease. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology. 2002;89(5):474-8.
26. Galanter WL, Hier DB, Jao C, Sarne D. Computerized physician order entry of medications and clinical decision support can improve problem list documentation compliance. Int J Med Inform. 2010;79(5):332-8.
27. Seidling HM, Phansalkar S, Seger DL, Paterno MD, Shaykevich S, Haefeli WE, et al. Factors influencing alert acceptance: a novel approach for predicting the success of clinical decision support. J Am Med Inform Assoc. 2011;18(4):479-84.
28. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc. 2006;13(2):138-47.
29. Phansalkar S, Desai AA, Bell D, Yoshida E, Doole J, Czochanski M, et al. High-priority drug-drug interactions for use in electronic health records. J Am Med Inform Assoc. 2012;19(5):735-43.

Table I. Allergist/immunologist involvement in undiagnosed and diagnosed clinical AERD. The charts of undiagnosed (n= 62) and diagnosed AERD (n= 438) cases were assessed for involvement by allergy/immunology specialists.

	AERD		
	Diagnosed	**Undiagnosed**	**Total**
Allergist/immunologist Involvement	408	24	
No Allergist/immunologist Involvement	30	38	
Total, n	438	62	500
Allergy Involvement %	93.2	38.7	

Table II. Demographics of diagnosed and undiagnosed AERD cases and the Brigham and Women's Hospital AERD registry. Statistical analyses run between diagnosed and undiagnosed AERD. The BWH AERD registry demographics have been included for reference. n – sample size; IQR – interquartile range; # - Fisher's exact test; ^ - Mann-Whitney U test; * - T test.

	Diagnosed AERD	**Undiagnosed AERD**	**p-value**	**AERD Registry**
Sample size, n	**438**	**62**		**96**
Male, n (%)[#]	**179 (40.9)**	**26 (41.9)**	**0.9**	**42 (43.8)**
Median age, years (IQR)^	**54 (45-65)**	**58 (51-72)**	**<.01**	**(42-52 60)**
Race, n (%)[#]			**0.7**	
White/Caucasian	**356 (81.3)**	**53 (85.5)**		**87 (90.6)**
Black/African American	**27 (6.2)**	**2 (3.2)**		**3 (3.1)**
Hispanic/Latino	**16 (3.7)**	**3 (4.8)**		**2 (2.1)**
Asian	**5 (1.1)**	**1 (1.6)**		**3 (3.1)**
Other/Unknown	**34 (7.8)**	**3 (4.8)**		**1 (1.0)**
Encounters, total, median (IQR)*	**37.5 (11-101)**	**54.5 (19-126)**	**0.3**	

Figure 1. Venn diagram of the clinical characteristics of cases identified by an AERD bioinformatics algorithm. From 2,647,842 patients seen within the Partners Healthcare system between 12/2004 and 11/2014 aged 18 and older, we identified cases with a diagnosis of asthma, nasal polyps, and/or NSAID allergy. NSAID allergy was restricted to only those with a specified respiratory reaction to NSAIDs or an unspecified ("unknown") reaction. The cohort of "possible AERD" cases, in yellow, lies at the intersection of all three clinical characteristics. n – sample size.

Figure 2. Flow chart for the assessment of the possible AERD cohort. PPV for identifying AERD in subjects with asthma, nasal polyposis and a recorded respiratory reaction to an NSAID (Cohorts A) = 88.7%. PPV for identifying AERD subjects having a recorded respiratory or unspecified reaction to an NSAID (Cohort A+B) not previously enrolled in the AERD registry = 78.4%. PPV for algorithm identifying all patients with AERD (Cohort A+B+AERD registry) = 81.1%. n – sample size. * - 732 charts were initially identified by the algorithm and one test chart was excluded.

Figure 3. Venn diagram of the possible AERD cases identified by the AERD algorithm (SQL#1-3), AERD specific search terms (SQL#4), and the BWH AERD Registry.

Figure 1. Venn diagram of the clinical characteristics of cases identified by an AERD bioinformatics algorithm.

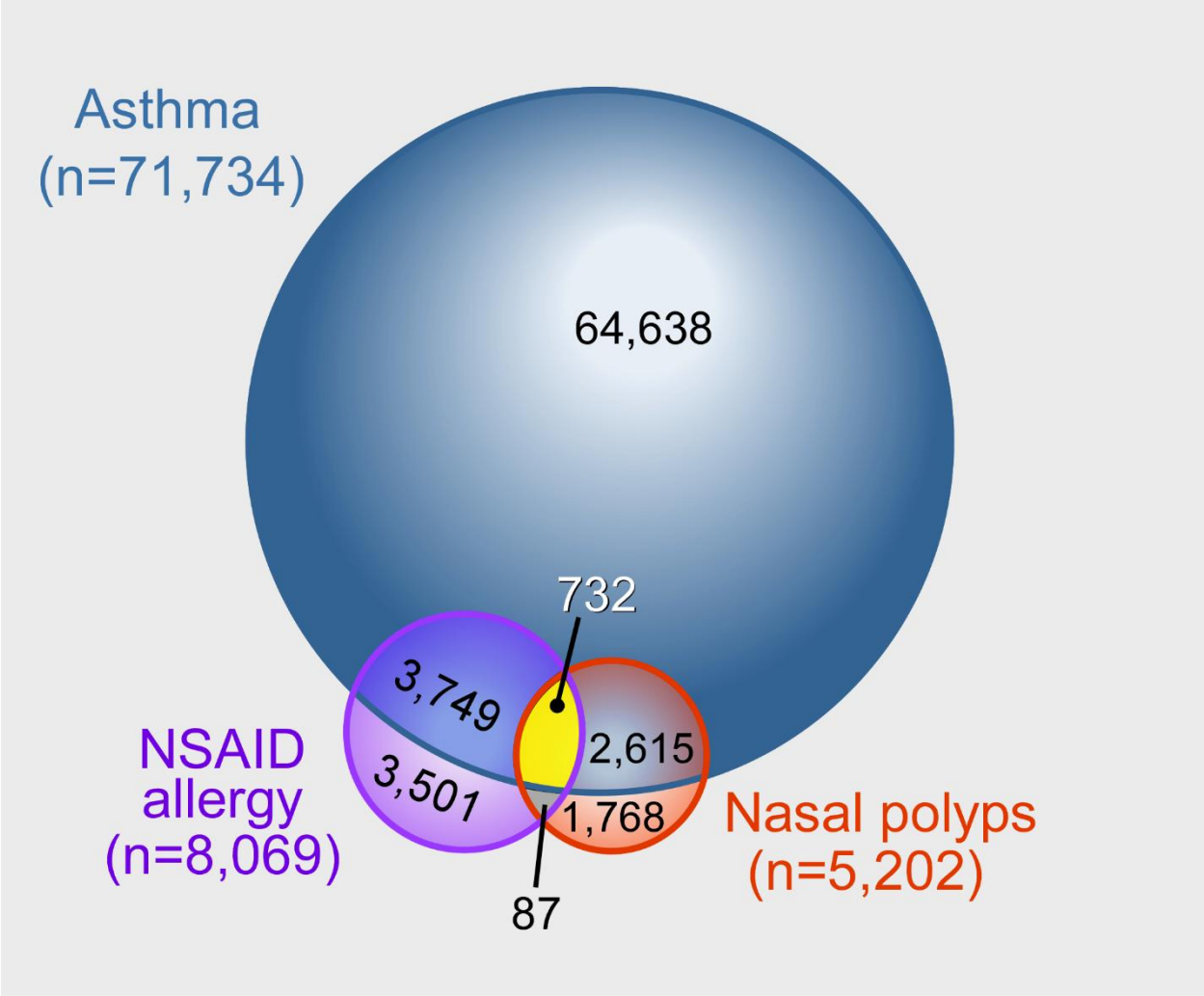


Figure 2. Flow chart for the assessment of the possible AERD cohort.

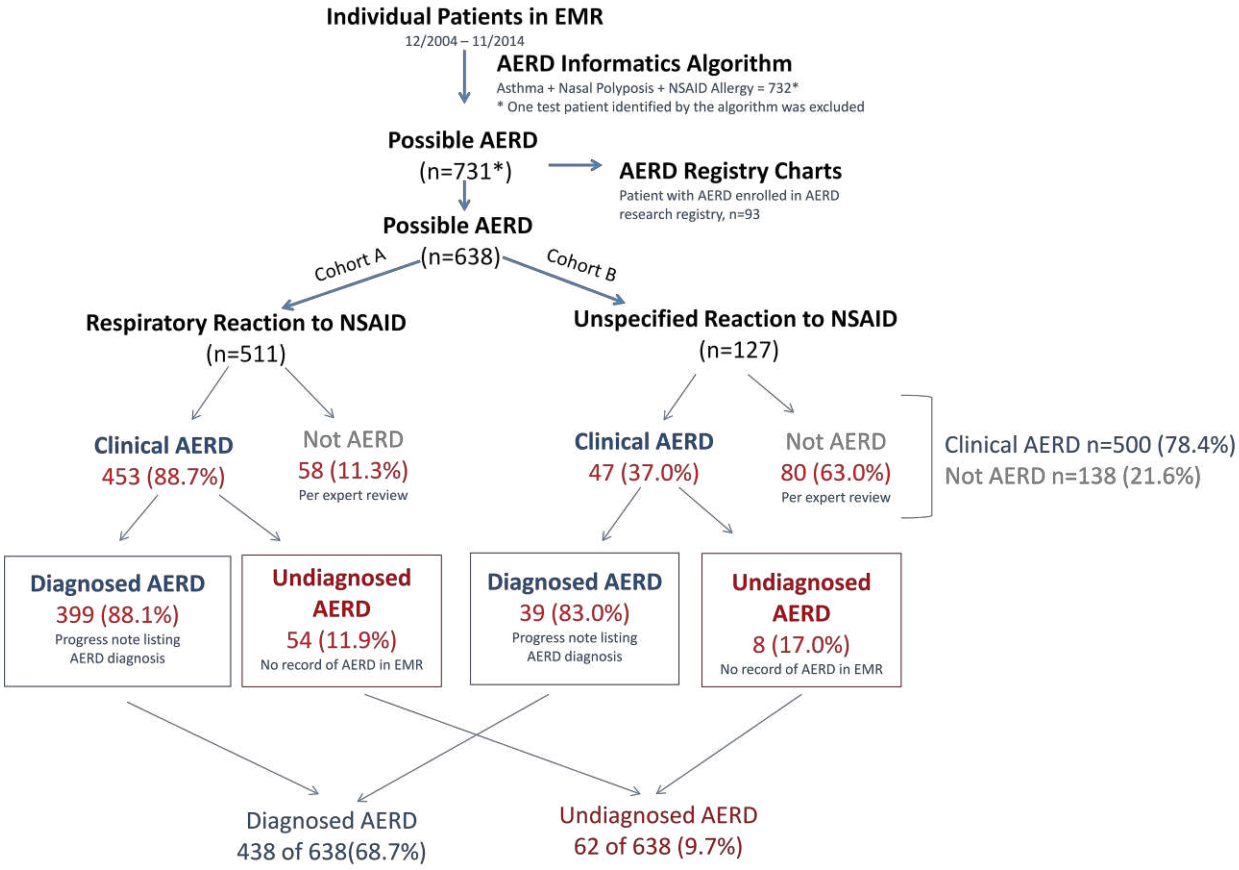
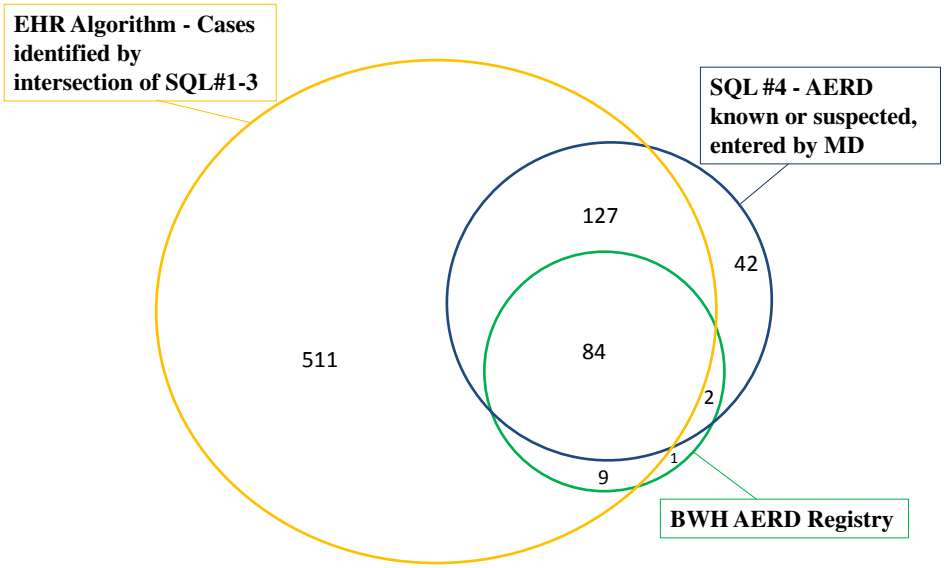


Figure 3. Venn diagram of the possible AERD cases identified by the AERD algorithm (SQL#1-3), AERD specific search terms (SQL#4), and the BWH AERD Registry.



Repository Legends:

Repository Table E1. Partners Research Patient Data Repository (RPDR) queries used to identify charts containing information supporting a diagnosis of asthma/nasal polyps and aspirin/non-steroidal anti-inflammatory drug (NSAID) allergy.

Repository Figure E1. Structured query language (SQL) query #1 developed to identify a patient population with asthma.

Repository Figure E2. Structured query language (SQL) query #2 developed to identify a patient population with nasal polyps.

Repository Figure E3. Structured query language (SQL) query #3 developed to identify a patient population with non-steroidal anti-inflammatory drug (NSAID) hypersensitivity reactions typical of the respiratory symptoms triggered by NSAIDs in aspirin-exacerbated respiratory disease (AERD) or charts that reported unspecified ("unknown") reactions to NSAIDs . Asa – aspirin. Sob – shortness of breath.

Repository Figure E4. Structured query language (SQL) query #4 developed to identify a patient population with known or suspected aspirin-exacerbated respiratory disease (AERD). Asa – aspirin. Sob – shortness of breath.

Repository Table 1:

Asthma/Nasal polyps RPDR query:

Wheezing 786.07-.09 OR
Asthma, all types and exacerbation states – 493.0-99 OR
Diagnosis-Related Groups for bronchitis and asthma age >17 OR
Prescription, inpatient or outpatient, in all forms:
- Albuterol– inhaler and nebulizer OR
- Ipratropium plus albuterol OR
- Levalbuterol OR
- Zileuton OR
- Budesonide and all other inhaled corticosteroids (ICS) in all formulation including
- ICS/Long-acting beta-agonist combo OR
- Montelukast OR
- Zafirlukast

AND

Anosmia 781.1 OR
Chronic rhinitis 472.0 OR
Nasal polyp 471 OR
Prescriptions for any nasal steroid – generic and brand name OR
Any procedure code for polypectomy - CPT 31288/30110, P2252/2264 OR
Any procedure code for nasal endoscopy - CPT31231

Aspirin/NSAID allergy RPDR query:

Personal history of aspirin allergy – V14.6 code OR
Desensitization – V071.XX code OR
Drug allergy NOS – 995.3 OR
Adverse effect of drug 995.27, 995.29 OR
Anaphylactic shock NOS 995.0 OR
Peripheral blood eosinophil count >500/µl OR
Eosinophilia 288.3

Repository Figure E1.

```
-- ASTHMA
select distinct patient_id
from (
select patient_id
from diagnoses
where (
diagnosis like "*bronchitis and asthma age >17*" or
diagnosis like "*asthma, unspecified without mention of status asthmaticus*" or
diagnosis like "*extrinsic asthma without mention of status asthmaticus*" or
diagnosis like "*asthma, unspecified type, with acute exacerbation*" or
diagnosis like "*extrinsic asthma with acute exacerbation*" or
diagnosis like "*chronic obstructive asthma, without mention of status asthmaticus*"
or
diagnosis like "*intrinsic asthma without mention of status asthmaticus*" or
diagnosis like "*extrinsic asthma with status asthmaticus*" or
diagnosis like "*chronic obstructive asthma with acute exacerbation*" or
diagnosis like "*asthma, unspecified type, with status asthmaticus*" or
diagnosis like "*intrinsic asthma, with acute exacerbation*" or
diagnosis like "*cough variant asthma*" or
diagnosis like "*intrinsic asthma with status asthmaticus*" or
diagnosis like "*chronic obstructive asthma, with status asthmaticus*" or
diagnosis like "*asthma, unspecified*" or
diagnosis like "*asthma*" or
diagnosis like "*extrinsic asthma*" or
diagnosis like "*chronic obstructive asthma*" or
diagnosis like "*asthma-lmr 29*" or
diagnosis like "*asthmatic bronchitis-lmr 30*" or
diagnosis like "*exercise-induced asthma-lmr 1586*" or
diagnosis like "*asthma, acute exacerbation-lmr 1288*" or
diagnosis like "*asthma-oncall*" or
diagnosis like "*asthmatic bronchitis-oncall*" or
diagnosis like "*exercise induced asthma*" or
diagnosis like "*exercise induced bronchospasm*"
) and not (
diagnosis like "*bronchitis and asthma age 0-17*" or
diagnosis like "*family history of asthma*" or
diagnosis like "*antiasthmatics causing adverse effects in therapeutic use*" or
diagnosis like "*asthma care model patient-oncall*"
)
union
select patient_id
from problems
where (
problem = "asthma" or
problem = "h/o asthma" or
problem = "allergic asthma" or
problem = "cough variant asthma" or
problem = "asthma - resolved" or
problem = "asthma, acute exacerbation" or
problem = "asthma/allergic rhinitis" or
problem = "moderate persistent asthma" or
problem = "severe persistent asthma" or
problem = "asthmatic breathing" or
problem = "extrinsic asthma" or
problem = "asthma - or eosinophilic bronchitis" or
problem = "asthma,severe" or
problem = "chronic obstructive asthma" or
problem like "*asthma, aspirin sensitive*" or
problem like "*asthma, frequent steroids*" or
problem like "*asthma, intubated*"
))
```

Repository Figure E2.

```
-- NASAL POLYPS
select distinct patient_id
from (
select patient_id
from diagnoses
where (
diagnosis like "*polyp of nasal cavity*" or
diagnosis like "*nasal polyp*" or
diagnosis like "*other polyp of sinus*" or
diagnosis like "*polypoid sinus degeneration*" or
diagnosis like "*sinus surgery, polyp*" or
diagnosis like "*sinus polyp*"
)
union
select patient_id
from problems
where (
problem like "*polyp of nasal cavity*" or
problem like "*nasal polyp*" or
problem like "*other polyp of sinus*" or
problem like "*polypoid sinus degeneration*" or
problem like "*sinus surgery, polyp*" or
problem like "*sinus polyp*"
or ((
problem like "*sinus*" or
problem like "*nasal*" or
problem like "*allergic rhinitis*"
) and (
comments like "*polyp*"
))
)
)
```

Repository Figure E3.

```
-- NSAID HYPERSENSITIVITY
select distinct patient_id
from allergies
where (
allergen like "*aspirin*" or
allergen = "asa" or
allergen like "* asa *" or
allergen like "*+asa *" or
allergen like "asa *" or
allergen like "*+asa+*" or
allergen like "asa-*" or
allergen like "* asa,*" or
allergen like "* asa" or
allergen like "asa,*" or
allergen like "asa/*" or
allergen like "*/asa/*" or
allergen like "* asa,*" or
allergen like "*,asa,*" or
allergen like "*nsaid*" or
allergen like "*ibuprofen*" or
allergen like "*ibuprophen*" or
allergen like "*advil*" or
allergen like "*motrin*" or
allergen like "*naproxen*" or
allergen like "*naprosyn*" or
allergen like "*indomethacin*" or
allergen like "*ketorolac*" or
allergen like "*toradol*" or
allergen like "*salicylic acid*" or
allergen like "*sulfasalazin*" or
allergen like "*olsalazin*" or
allergen like "*sulindac*" or
allergen like "*etodolac*" or
allergen like "*flurbiprofen*" or
allergen like "*ketoprofen*" or
allergen like "*fenoprofen*" or
allergen like "*oxaprozin*" or
allergen like "*mefenamic acid*" or
allergen like "*meclofenamic acid*" or
allergen like "*piroxicam*" or
allergen like "*meloxicam*" or
allergen like "*diclofenac*"
) and (
reaction like "*bronchospasm*" or
reaction like "*brochospasm*" or
reaction like "*bronchoconstriction*" or
reaction like "*shortness of breath*" or
reaction like "*sob*" or
reaction like "*chest tightnes*" or
reaction like "*asthma*" or
reaction like "*ashtma*" or
reaction like "*anaphyla*" or
reaction like "*amaphyla*" or
reaction like "*anaphylla*" or
```

```
reaction like "*anaphlaxis*" or
reaction like "*cough*" or
reaction like "*wheez*" or
reaction like "*nasal polyp*" or
reaction like "*nasla polyp*" or
reaction like "*nasalpolyp*" or
reaction like "*asthma, polyp*" or
reaction like "*nasal stuffines*" or
reaction like "*nasal congestion*" or
reaction like "*congestion/nasal*" or
reaction like "*develops polyps*" or
reaction like "*rash*" or
reaction like "*flushing*" or
reaction like "*sneezing*" or
reaction like "*resp. react*" or
reaction like "*respiratory distres*" or
reaction like "*unable to breath*" or
reaction like "*difficulty breathing*" or
reaction like "*difficult to breath*" or
reaction like "*trouble breathing*" or
reaction like "*aerd*" or
reaction like "*sampter*" or
reaction like "*santer*" or
reaction like "*samter*" or
reaction like "*exacerbated respiratory disease*" or
reaction like "*unknown*"
)
```

Repository Figure E4.

```
-- KNOWN OR SUSPECTED AERD
select distinct patient_id
from (
select patient_id, problem as feature
from problems
union
select patient_id, comments as feature
from problems
union
select patient_id, problem_code_description as feature
from problems
union
select patient_id, allergen as feature
from allergies
union
select patient_id, reaction as feature
from allergies
)
where (
feature like "*aerd*" or
feature like "*aspirin-induced asthma*" or
feature like "*aspirin induced asthma*" or
feature like "*aspirin-induced respiratory*" or
feature like "*aspirin induced respiratory*" or
feature like "*aspirin exacerbated respiratory*" or
feature like "*aspirin-exacerbated respiratory*" or
feature like "*exacerbated respiratory disease*" or
feature like "*aspirin-sensitive asthma*" or
feature like "*aspirin sensitive asthma*" or
feature like "*aspirin causes shortness of breath*" or
feature like "*aspirin causes sob*" or
feature like "*nsaids, bronchospasm or wheezing*" or
feature like "*nsaid- breathing difficulty/bronchospasm*" or
feature like "*samter*" or
feature like "*sampter*" or
feature like "*santer*" or
feature like "*triad asthma*" or
feature like "*motrin, ibuprofen in high doses over a prolonged
periodbronchospasm, wheezing*" or
feature like "*tartrazine (yellow dye#5) - anaphylaxis, asa - asthma*" or
feature like "*intolerant to asa as it worsens her asthma symptoms*" or
feature like "*avoids nsaids because of effect on asthma*" or
feature like "*aspirin cuz asthma attack*" or
feature like "*asa and nsaids cause hives and sob*" or
feature like "*asthma*nasal polyp*intoleran*nsaid*" or
feature like "*asa sensitivity and nasal polyp*" or
feature like "*asa-sensitivity and nasal polyp*" or
feature like "*aspirin allergy*nasal*polyp*" or
feature like "*aspirin-allergy*nasal*polyp*" or
feature like "*aspirin sensitivity*nasal*polyp*" or
feature like "*aspirin-sensitivity*nasal*polyp*" or
feature like "*asa allergy*nasal*polyp*" or
feature like "*asa-allergy*nasal*polyp*" or
feature like "*motrine and tylenol gets sob*")
```